



## 2012 International Conference on Applied Physics and Industrial Engineering An Ant Colony Clustering Algorithm Improved from ATTA

Jinbiao Wang, Ailing Tu, Hongwei Huang

*School of Computer Science and Technology,  
Civil Aviation University of China  
TianJin, China, 300300*

---

### Abstract

DENNEUBOURG presents the first ant-based clustering algorithm in 1991. Ant colony clustering has the characteristics of automatically clustering, high clustering accuracy, irregular cluster shapes and so on. These are important for a clustering algorithm, so ant colony clustering is arousing more and more data mining researchers' concentration. In 2004, J. Handl together with his partners normalized the basic ant colony clustering algorithm, getting the ATTA model with superior performance. In this paper, we'll present an algorithm improved from ATTA which we call it LCA and test LCA on UCI data sets to show its feasibility.

© 2011 Published by Elsevier B.V. Selection and/or peer-review under responsibility of ICAPIE Organization Committee. Open access under [CC BY-NC-ND license](#).

*Keywords:* Ant colony clustering; Logic; cold; UCI data sets; feature attribute

---

### 1. Atta Model

In the past two decades, many achievements in computer science depends on the bio-inspired information. Such as ant populations of parallelism, self-organization and robustness that played a particularly important role in promoting the development of intelligent algorithms. Although it is composed of simple individuals, but the entire population in case of no central control was able to complete complex tasks. Deneubourg J L in 1991 was enlightened by the behavior of ant colony classified ant eggs, first proposed Ant Colony Clustering Algorithm[1]. The algorithm was among the first to be used for the robot. Robot to imitate the behavior of ants, which is to identify two different objects, according to local environmental information to decide to pick up or drop the object, and so forth to achieve the purpose of object clustering. Then in 1994, Lumer and Fata used the idea of ant colony clustering to data analysis, proposed the LF algorithm<sup>[2]</sup>, of ant colony clustering mainly in three aspects: 1) perception radius, 2) pick up and drop probability calculation, 3) mobile mechanism. In 2004, J. Handl and others worked on the specification of the ant colony clustering algorithm such as the norm, got the superior performance ATTA model<sup>[3]</sup>.

Compared with the LF, ATTA's similarity function, picking up and dropping function, perception radius, and its memory parameters are all more normalized:

### 1.1. Three principal elements

①The similarity function:

$$f^*(i) = \begin{cases} \frac{1}{\sigma^2} \sum_{j \in L} (1 - \frac{\delta(i,j)}{\alpha}), & \text{if } \forall j (1 - \frac{\delta(i,j)}{\alpha}) > 0 \\ 0, & \text{else} \end{cases} \quad (1)$$

②Individual ants picking up and dropping probability formula :

$$p_{pick}^*(i) = \begin{cases} 1, & \text{if } f^*(i) \leq 1 \\ \frac{1}{f^*(i)^2}, & \text{else} \end{cases} \quad (2)$$

$$p_{drop}^*(i) = \begin{cases} 1, & \text{if } f^*(i) \geq 1 \\ f^*(i)^4, & \text{otherwise} \end{cases} \quad (3)$$

These two formulas derived from Experiments ,and the two formulas achieve better results accompanied by a perceived increase of radius.

③Perception radius  $r$  : During the operation of the algorithm, so that the radius of perception changed all over the time, ATTA model for linear growth of  $r$  from 1 to 5, when  $r$  changes,  $\sigma = 2r + 1$  also changes, because of  $r = (\sigma - 1)/2$ . But here the data object in order to maintain sufficient separation in space, so when  $r$  changes,  $\sigma$  constant.

### 1.2. Several important parameters

①Memory  $m$  : Similarity  $f^*(i)$  rather than dissimilarity  $\delta(i,j)$  used as a criterion, when the ants pick up a new individual data  $I$ , it put the data into memory in the neighborhood of the point position, calculating the similarity  $f^*(i)$  of those  $m$  locations. The most appropriate location of  $f^*(i)$  for the largest object in the down position of the data.

② $\alpha$  value changes over time : Initially given for each ant a unified single value  $\alpha$  between 0 to 1, after that, according to each ant's mobile data object to get change. That ant steps in the mobile

$\#active=100$ , the calculation of the data object that controls the failure rate  $r_{fail} = \frac{\#fail}{\#active}$ ,  $\#fail$  is the number of ants failing to put down its data objects,  $\#active$  is the number of ant movements. there is a new formula of  $\alpha$  :

$$\alpha \leftarrow \begin{cases} \alpha + 0.01, & \text{if } r_{fail} > 0.99 \\ \alpha - 0.01, & \text{if } r_{fail} \leq 0.99 \end{cases}$$

③Empirical parameters :

Number of ants =10,  
Memory  $m=10$ ,  
 $tstart=0.45*\#iterations$ ,

tend=0.55\*#iterations

space =  $\sqrt{10 \times \#items} \times \sqrt{10 \times \#items}$

stepsize =  $\sqrt{20 \times \#items}$  #iterations = 2000×#items

#iterations than or equal to 100 million.

ATTA clustering performance has been achieved relatively good condition.

However, For classes and extract other issues cluster adhesion problems that exist in the basic model, ATTA has not fully solved; Secondly, ATTA clustering accuracy and speed to be further improved.

ATTA has an interesting feature: after the formation of the initial class clusters, add an episode stage, that is, between the tstart and tend to change the similarity function  $f^*(i)$ , using  $1/N_{occ}$  instead of  $1/\sigma^2$ , here  $N_{occ}$  is the number of grid occupied by data objects in the local neighborhood. In this episode, the data points go through the "divergence - to re-aggregation" process, which improves the quality of clustering.

## 2. LCA: An improved ant colony clustering algorithm

LCA(Logic based cold ants) makes the following improvements on ATTA : Ant populations were initially picked up the data object, and calculate their current location is suitable for down, take the data object which are carried by those ants are not suitable for put down directly to the various objects for maximum similarity value of the position, moreover, in order to allow rapid formation of class cluster center, we propose a similarity measure based on logic. That means that : ant makes objects into similar and dissimilar two categories, and enthusiasm of similar objects (attract), not similar object detached (not processed).

### 2.1 Similarity calculation based on logic

ATTA seek all the properties of the data object to calculate the precise degree of similarity neighborhood (see formula(1)) ,but the reality is that many attributes are mutually weaken , and noises component contained in the data objects are also a wide gap, Therefore, accurate calculation often become wishful thinking. Practice by a large number of calculations found that the average distance  $\alpha$  of the data object is a key parameter. Less than  $\alpha$  similar, otherwise dissimilar. As shown below, data object  $i$  is carried by an ant at the center of the grid, Assuming  $\alpha$  is 0.2, data object  $\delta(i,j)$  around the neighborhood 8 grid is shown in Table 1.

1 △	8 ○	7 
2 ○	i ○	6 ○
3 ○	4 ○	5 △

Figure 1

TABLE 1

j	1	2	3	4	5	6	7	8
$\delta(i, j)$	0.6	0.1	0.15	0.12	0.4	0.08	null	0.1

Similarity function formula based on the ATTA (1) :

$$\begin{aligned}
 f^*(i) &= \frac{1}{\sigma^2} \sum_{j \in L} \left(1 - \frac{\delta(i, j)}{\alpha}\right) \\
 &= \frac{1}{\sigma^2} \left[ \left(1 - \frac{0.6}{0.2}\right) + \left(1 - \frac{0.1}{0.2}\right) + \left(1 - \frac{0.15}{0.2}\right) \right. \\
 &\quad \left. + \left(1 - \frac{0.12}{0.2}\right) + \left(1 - \frac{0.4}{0.2}\right) + \left(1 - \frac{0.08}{0.2}\right) + \left(1 - \frac{0.1}{0.2}\right) \right] \\
 &= \frac{1}{\sigma^2} (-0.75) \\
 &= 0
 \end{aligned}$$

Calculated result is that the probability that ant puts the data into that region is 0. Obviously, this result is wrong. The reason is simply because the two different points of the  $\delta(i, j)$  value is too large. Adjusted as follows for this :

- ① Statistics the number of similar points (snum) and the number of dissimilar points (dnum) in the neighborhood, if  $\left(1 - \frac{\delta(i, j)}{\alpha}\right) > 0$ ,
- ② that the data object i and j are alike, snum plus 1. The method for calculation of similarity to :

$$f'(i) = \begin{cases} 0, & \text{if } dnum > snum \\ snum, & \text{else} \end{cases} \quad (4)$$

In the example point 2, point 3, point 4, point 6 and point 8 are similar, snum=5; point 1 and point 5 are not similar, dnum=2. snum>dnum. The result is that ants lay down the data in this location, in line with the actual situation.

- ② Drop probability by :

$$p_{drop}^*(i) = \begin{cases} 1, & \text{if } f^*(i) > 4 \\ \left(\frac{f^*(i)}{4}\right)^3, & \text{otherwise} \end{cases} \quad (5)$$

The formula is obtained by experiment. The algorithm does not need to pick up the object probability formula.

③ Perception radius r : Throughout the algorithm the perception of the radius of the ant is constantly changing, to allow more data to be a part of the single action of ants. The Perception Radius of the proposed algorithm have linear increased from 1 to 3, and then decreased from 3 to 1.

④ Separation of the initial class of cluster centers: The latter part of the algorithm may occur two Class cluster adhesion object suitable for being put down , and then choose one of which to mark as not appropriate to lay down .

## 2.2 Algorithm flow

// Initialization phase

① The data is randomly scattered on a two-dimensional toroidal plane ;

② For(j=1; j<= number of ants; j++)

{

```

The ant j randomly selected a free data object and pick it up ;
The ant j randomly selected a location of the environment and jump to the empty position ;
}
// The main loop phase
③for(iteration=1;iteration<= the total number of iterations; iteration++)
{
    Calculate all the labels of ants which are suitable to put down their data objects;
    Increase the separation of the initial cluster centers to change some suitable places into non-down positions;
    The ants which not suitable for putting down jump into those suitable for putting down, in the neighborhood of the
maximum similarity value;
    Update the free data objects table;
    If one ant successfully put down the object
    {
        Randomly finds a new free data object I;
        Ant picks up the new data object i;
        Ant jumps onto the location of the object I;
    }
}
}

```

### 2.3 Simulation of the algorithm

①400 data objects, 10 ants, ATTA model 100 times to run the process and results:

Iteration=0:      Iteration=70:      Iteration=100:



Figure 2. clustering process and results for ATTA algorithms

In the model ATTA, Previous algorithm requires a longer period of time to run. After forming in the class cluster, the algorithm used the stage of an episode. At this stage, ants consider only the similarity without considering the density to pick up and down the object data, as a results, data objects will be orderly dispersed. When the data points are ordered dispersed, then use of previous similarity formula to Cluster of data objects, this process is equivalent to re-select the cluster center of class. Because the ant colony clustering algorithm is a stochastic algorithm, class of the results of cluster center re-selection has the potential to separate from each other class of cluster center, may also make the new class of cluster center still closely linked. The above model, the lower left corner of the class of cluster centers were separated, but separation was not obvious.

②400 data objects, 100 ants, LCA model 1 time to run the process and results.

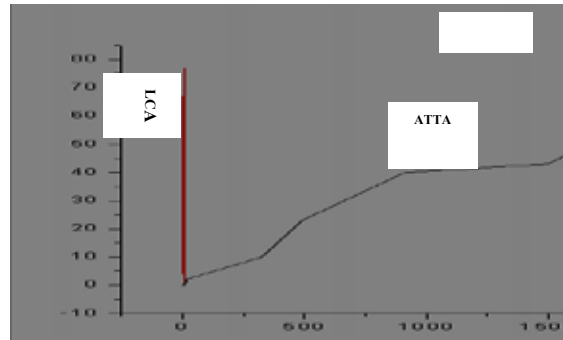
step=0:      step =7:      step =19:



Figure 3. Run Results of LCA

Figure 3 shows that LCA only use 19 steps to a good accurate classes cluster together. After the first step of the algorithm, cluster centers have been formed in all categories. Careful observation can be found, the early formation of the lower right corner there is a cluster of red type, but the center is too close to the blue and purple cluster space, so it will not be put down as a suitable location. As the algorithm progresses, the red class cluster clustering slowly up to that position that far apart with others .

- ③ LCA calculations faster than ATTA increased by 1 ~ 2 orders of magnitude. That is because in the clustering process, a huge gap between the number of data move caused.



**Figure 4.** ATTA and the LCA algorithm to cluster the results of each comparison the number of moving objects

In Figure 4, X axis represents the number to be moving, Y-axis is the number corresponding to the number of objects to move. The average number of ATTA move which is 2129 times while the number of LCA is 5 times.

### 3. Experiments on UCI Data sets

Examples in Table 2 are from the UCI data sets <sup>[4]</sup>.

An object is often expressed through the multidimensional attribute, but clustering practice shows that, only feature attribute contribute to clustering. The following data set of wine as an example for analysis. To get feature attributes, we first experiment the 13-dimensional properties of single-attribute of wine in the algorithm experimental platform, Through a combination of 7 and 13, the clustering results are good, we may conclude that 7 and 13 are feature attributes.

Through UCI data sets typical examples of the clustering effect of comparison, we can see the LCA algorithm have significant improvements than ATTA and the classic LF. As indicated in Table 2.

attribute 7                  attribute 13                  attributes 7,13



**Figure 5** Clustering result

TABLE 2 clustering results of data set using different algorithms

Data Set	Iris			Wine			Soybean			Lenses			Balances			Ecoli		
Evaluate Index	LF	ATTA	LCA	LF	ATTA	LCA	LF	ATTA	LCA	LF	ATTA	LCA	LF	ATTA	LCA	LF	ATTA	LCA
MacroP(%)	50	72	88	48	61	80	43	59	100	47	44	58	62	79	55	26	27	32
MacroR(%)	89	82	90	88	84	88	69	61	100	85	76	93	58	12	85	81	81	79
MacroF1(%)	63	76	89	61	69	83	53	59	100	59	55	72	58	21	66	39	40	46
C(%)	63	43	100	66	66	100	64	65	100	55	55	66	-136	-1299	79	54	55	49
Iterative rounds	70	150	1	70	180	1	80	210	1	50	100	1	70	250	1	70	250	1
Number of steps per round	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>2</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>2</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>2</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>2</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>2</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>2</sup>
Iteration Time (s)	55	69	3.8	106	88	4.9	59	85	3	79	37	2.7	311	162	32	56	132	9.5
Size example	150			178			47			24			625			336		
Dimension	4			13			35			4			4			7		

Explanation :

①MacroP is the macro average accuracy :

$$MacroP = \frac{1}{n} \sum_{i=1}^n \frac{I_i}{M_i} \quad (6)$$

Thereinto,  $M_i$  is the number of clustering elements in  $i$ -class,  $I_i$  is the correct clustering number of elements in  $M_i$ .  $n$  is the total number of all categories.

②MacroR is the macro average recall rate :

$$MacroR = \frac{1}{n} \sum_{i=1}^n \frac{I_i}{N_i} \quad (7)$$

Thereinto,  $N_i$  is the number of elements in  $i$ -class,  $I_i$  is the number of elements clustered in  $i$ -class.  $n$  is the total number of all categories.

③MacroF1 is the average F1 value of the macro :

$$MacroF1 = \frac{MacroP \times MacroR \times 2}{MacroP + MacroR} \quad (8)$$

Thereinto, MacroP is the macro average accuracy, MacroR is the macro average recall-rate.

④Clustering data rate:

$$C = \begin{cases} \frac{n1}{n2}, & \text{if } n1 \leq n2 \\ -\frac{n1-n2}{n2}, & \text{else} \end{cases} \quad (9)$$

$C$  more closer to 1 indicates better clustering results. Thereinto,  $n1$  means that clustering system actually produce the number of polymer,  $n2$  means that the classes existed in real data.

#### 4. Summary and Discussion

By this paper proposed, LCA algorithm is useful and feasible. Examples show that ants mode Environmental Recognition Judgement Method logical worth more than accurate calculations in depth. An actual problem is not only facing the multidimensional, but inevitably existing all kinds of noise, therefore, how to more effectively identify the characteristics of property, move the noise is the direction of future efforts. Especially with the advent of multi-core computer the ant algorithm is possible to achieve parallel computation , so that prospect of ant clustering algorithm is worth the wait.

#### 5. Acknowledgment

This research is supported by the National Natural Science Foundation of China (Grant Nos. 60472121 and 60979021).

The authors are grateful to the editors and the anonymous reviewers: their remarks and suggestions are important for shaping this paper.

#### References

- [1]J.L.Deneubourg, S.Gross, N.R.Franks, A.Sendova-Franks, C.Detrain, and L.Chretien. The dynamics of collective sorting: Robot-like ants and ant-like robots. *Simulation of Adaptive Behavior: From Animals to Animats*, pp.356-363, 1991.
- [2]E.D.Lumer and B.Faieta.Diversity and Adaptation in populations of clustering ants.In *Proc. Of the Third International Conference on the Simulation of Adaptive Behavior:From Animals to Animats 3*,pp.449-508.Mit Press,1994.
- [3]J.Handl, J.Knowles, and M.Dorigo. Ant-based clustering and topographic mapping. *Artificial Life*, Vol.12(1), 2004.
- [4]<http://archive.ics.uci.edu/ml/datasets.html>